WHAT IS CLAIMED IS:

1. A method for matching patterns in a string of symbols comprising:

identifying a first pattern of symbols to be matched, wherein the first pattern contains a prefix pattern, a value pattern and a suffix pattern;

identifying candidate matches for the first pattern in the string, wherein each candidate match for the first pattern includes a candidate match for the prefix pattern, a candidate match for the suffix pattern and a candidate match for the value pattern;

determining a cost associated with each of the candidate matches for the first pattern, wherein the cost associated with each of the candidate matches for the pattern includes a cost associated with the corresponding candidate match for the prefix pattern, a cost associated with the candidate match for the suffix pattern and a cost associated with the candidate match for the value pattern; and

selecting one or more candidate matches for the pattern that meet a cost selection criterion.

2. The method of claim 1 wherein determining a cost associated with each of the candidate matches comprises calculating a corresponding edit distance.

3. The method of claim 1 wherein identifying the first pattern comprises providing a single example string wherein the first pattern is selected from the example string.

4. The method of claim 1 further comprising examining the string to identify spans of interest, wherein each of the spans of interest meets a specified filtering criterion.

5. The method of claim 4 wherein the specified filtering criterion comprises the inclusion of a keyword.

6. The method of claim 1 wherein selecting one or more candidate matches for the pattern that meet a cost selection criterion comprises selecting one or more candidate matches that have corresponding costs which fall below a selected threshold.

7. The method of claim 1 wherein selecting one or more candidate matches for the pattern that meet a cost selection criterion comprises selecting a predetermined number of candidate matches that have the lowest corresponding costs.

8. The method of claim 1 wherein selecting one or more candidate matches for the pattern that meet a cost selection criterion comprises selecting a candidate match that has a lowest cost and selecting additional candidate matches that have corresponding costs which are within a predetermined tolerance of the lowest cost.

9. The method of claim 1 further comprising adjusting the cost selection criterion and selecting one or more candidate matches for the pattern that meet the adjusted cost selection criterion.

10. The method of claim 1 wherein the cost associated with the corresponding candidate match for the prefix pattern, and the cost associated with the candidate match for the suffix pattern are more heavily weighted than the cost associated with the candidate match for the value pattern.

11. The method of claim 1 wherein the cost associated with each of the candidate matches for the first pattern is determined by adding the cost associated with the corresponding candidate match for the prefix pattern, the cost associated with the candidate match for the suffix pattern and the cost associated with the candidate match for the value pattern.

12. The method of claim 1 wherein identifying each candidate match for the first pattern comprises identifying the candidate match for the prefix pattern, wherein the candidate match for the prefix pattern defines a first end of a value window, then identifying a corresponding candidate match for the suffix pattern, wherein the candidate match for the suffix pattern defines a corresponding second end of the value window, wherein the candidate match for the value pattern comprises the symbols within the value window.

13. The method of claim 1 further comprising filtering the candidate match for the value pattern using a keyword.

14. The method of claim 1 further comprising filtering the candidate match for the value pattern using a regular expression.

15. The method of claim 1 wherein identifying candidate matches for the prefix pattern comprises constructing an edit distance matrix for the prefix pattern and identifying one or more candidate matches for the prefix pattern, constructing an edit distance matrix for the suffix pattern and identifying one or more candidate matches for the suffix pattern, and identifying a candidate match for the value pattern between each pair of candidate prefix matches and candidate suffix matches.

16. A computer readable medium containing instructions which are configured to implement the method comprising:

identifying a first pattern of symbols to be matched, wherein the first pattern contains a prefix pattern, a value pattern and a suffix pattern;

identifying candidate matches for the first pattern in the string, wherein each candidate match for the first pattern includes a candidate match for the prefix pattern, a candidate

match for the suffix pattern and a candidate match for the value pattern;

determining a cost associated with each of the candidate matches for the first pattern, wherein the cost associated with each of the candidate matches for the pattern includes a cost associated with the corresponding candidate match for the prefix pattern, a cost associated with the candidate match for the suffix pattern and a cost associated with the candidate match for the value pattern; and

selecting one or more candidate matches for the pattern that meet a cost selection criterion.

17. The computer readable medium of claim 16 wherein determining a cost associated with each of the candidate matches comprises calculating a corresponding edit distance.

18. The computer readable medium of claim 16 wherein identifying the first pattern comprises providing a single example string wherein the first pattern is selected from the example string.

19. The computer readable medium of claim 16 further comprising examining the string to identify spans of interest, wherein each of the spans of interest meets a specified filtering criterion.

20. The computer readable medium of claim 15 wherein the specified filtering criterion comprises the inclusion of a keyword.

21. The computer readable medium of claim 16 wherein selecting one or more candidate matches for the pattern that meet a cost selection criterion comprises selecting one or more candidate matches that have corresponding costs which fall below a selected threshold.

22. The computer readable medium of claim 16 wherein selecting one or more candidate matches for the pattern that meet a cost selection criterion comprises selecting a predetermined number of candidate matches that have the lowest corresponding costs.

23. The computer readable medium of claim 16 wherein selecting one or more candidate matches for the pattern that meet a cost selection criterion comprises selecting a candidate match that has a lowest cost and selecting additional candidate matches that have corresponding costs which are within a predetermined tolerance of the lowest cost.

24. The computer readable medium of claim 16 further comprising adjusting the cost selection criterion and selecting one or more candidate matches for the pattern that meet the adjusted cost selection criterion.

25. The computer readable medium of claim 16 wherein the cost associated with the corresponding candidate match for the prefix pattern, and the cost associated with the candidate match for the suffix pattern are more heavily weighted than the cost associated with the candidate match for the value pattern.

26. The computer readable medium of claim 16 wherein the cost associated with each of the candidate matches for the first pattern is determined by adding the cost associated with the corresponding candidate match for the prefix pattern, the cost associated with the candidate match for the suffix pattern and the cost associated with the candidate match for the value pattern.

27. The computer readable medium of claim 16 wherein identifying each candidate match for the first pattern comprises identifying the candidate match for the prefix

pattern, wherein the candidate match for the prefix pattern defines a first end of a value window, then identifying a corresponding candidate match for the suffix pattern, wherein the candidate match for the suffix pattern defines a corresponding second end of the value window, wherein the candidate match for the value pattern comprises the symbols within the value window.

28. The computer readable medium of claim 16 further comprising filtering the candidate match for the value pattern using a keyword.

29. The computer readable medium of claim 16 further comprising filtering the candidate match for the value pattern using a regular expression.

30. The computer readable medium of claim 16 wherein identifying candidate matches for the prefix pattern comprises constructing an edit distance matrix for the prefix pattern and identifying one or more candidate matches for the prefix pattern, constructing an edit distance matrix for the suffix pattern and identifying one or more candidate matches for the suffix pattern, and identifying a candidate match for the value pattern between each pair of candidate prefix matches and candidate suffix matches.